PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (http://bmjopen.bmj.com/site/about/resources/checklist.pdf) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	The prevalence of myopia and high myopia, and the association with education: Shanghai Child and Adolescent Large-scale Eye Study (SCALE), a cross-sectional study
AUTHORS	He, Xiangui; Sankaridurg, Padmaja; Xiong, Shuyu; Li, Wayne; Naduvilath, Thomas; Lin, Senlin; Weng, Rebecca; Lv, Minzhi; Ma, Yingyan; Lu, Lina; Wang, Jingjing; Zhao, Rong; Resnikoff, Serge; Zhu, Jianfeng; Zou, Haidong; Xu, Xun

VERSION 1 – REVIEW

REVIEWER	Pan, Chen-wei Medical College of Soochow University, School of Public Health
REVIEW RETURNED	08-Feb-2021

GENERAL COMMENTS	This paper is well-written and the method seems to be sound. The sample size is very large and across the various districts and ages. To the best of my knowledge, the sample size is the largest among previous published school-based studies. I have only some minor comments for the authors to consider. 1.The authors used a regression discontinuity model to better understand the effect of education. But the model is not a commonly used model in epi studies. The authors may want to provide more details of the model in the methods section. 2. When comparig the prevalence estimates of myopia in the discussion section, it is better to inlude a figure. In addition, the authors might include more studies reporting the prevalence of myopia in Chinese children. 3.The authors mentioned information bias using questinnaires. It is better if they could explain if the bias is differential or non-differential and its potiential impact on effect size of ORs.

REVIEWER	Baraas, R University of South-Eastern Norway, National Centre for Optics, Vision and Eye Care
REVIEW RETURNED	03-Mar-2021

GENERAL COMMENTS	I find the data presented in this manuscript very interesting. The dataset is large and there is no doubt that myopia is a major public health burden in Shanghai, both today and in the future. There are aspects that would benefit from more information and a more

	elaborate discussion, especially related to the causal effect of the school start date on the refractive error. The authors have not addressed important uncertainties that are needed to substantiate that the reported 0.67D difference over 10 years is “a striking effect of schooling/education resulting in a more myopic refractive error”? Two steps have been made by the authors for ensuring a conservative prevalence estimate of myopia when non-cycloplegic data is the basis for the classification. First, myopia was defined as =< -1.00 D. Second, a model (Equation 1) was used to correct for known errors with classification of refractive error based on non-cycloplegic autorefraction. Why did the authors choose -1.00D as the alternative definition rather than, let’s say, -0.75 or -1.25D? Was this related to the uncertainty with the model in reclassifying individuals or was it arbitrarily chosen to fit previously published prevalence estimates? The model reportedly only classifies 77% of the eyes correctly (Sankaridurg et al. 2017). What are the consequences that over 20% of eyes could have been misclassified, both in relation to prevalence, but also to the outcome of the regression discontinuity analysis? The model includes visual acuity measures. It is a strength that logMAR charts were employed, and that luminance was measured in the cabinets used, however, the range of luminance was considerable (80–320 cd/m2), could this have influenced the percentage of misclassifications as the model was based on visual acuity measures obtained at 400 cd/m2 (Sankaridurg et al. 2017)? Regression discontinuity design (RDD) is based on a number of assumptions, and an assessment of whether these assumptions are met in this study needs to be included. An advantage of RDD is its ability to estimate local causal effects in a small range near the cut-off, but there is no information on what the range was: was it all the children born in August versus all the children born in September? There are discontinuities at locations other than the cut-off: might these temporal trends confound the results? Only means are presented in Figure 1, there are no actual data points. Neither are the actual regression lines with confidence intervals on each side of the identified threshold included (in Figure 1). Thus, the novel aspect in this manuscript, regression discontinuity (RD) analysis and the strength of the results from this analysis is difficult to evaluated properly. Error bars are given in Figure 2, but there is no information on whether these are SDs, SEs or CIs. Another issue is the asymmetry in number of participants in the pre- versus post- 1 September groups (339 963 in the pre group compared with 175 181 in the post group): are more children born pre- 1 September, and if so to what degree could this have affected RD analysis and results? Are there differences in socio-economic status, birth rates and frequency of caesarean section pre- and post- 1 September across the 17 districts in Shanghai, and if so to what degree could this have affected RD analysis and results? The uncertainties related to the above-mentioned issues have not been addressed by the authors.

	Known issues related to data from questionnaires has been acknowledged by the authors. Other comments: Objectives in abstract refers to "geographical variation", but there are no results related to the geographical variation across the 17 districts in Shanghai. There are no error bars in Figure 3. Line 41: It is unclear what information was given to participants and if it was the participant or one or both parents/carers who gave consent. Was the questionnaire validated in some way? The questionnaire should be included as an appendix.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Dr. Chen-wei Pan, Medical College of Soochow University
Comments to the Author:
This paper is well-written and the method seems to be sound. The sample size is very large and across the various districts and ages. To the best of my knowledge, the sample size is the largest among previous published school-based studies. I have only some minor comments for the authors to consider.

1.The authors used a regression discontinuity model to better understand the effect of education. But the model is not a commonly used model in epi studies. The authors may want to provide more details of the model in the methods section.
We have amended the methods section as follows and also provided a reference.

The interrelationship between age, education and refractive error was evaluated using a regression discontinuity (RD) model. RD model is used to estimate the impact of a policy or program in situations where exposure to a risk factor is based on whether they exceed or fall behind a designated cut-off point. In the present analysis, we considered education as a risk factor. Children born in a given year (same age) were assigned to either pre or post-September groups based on the school entry cut-off criteria of 1 September; those born pre-September are admitted to a higher class/grade compared to those born on or post 1 September. Thus, the aim was to determine if for a given age, children born pre- September had a more myopic refractive error compared to post-September as they were in a higher class at school (greater academic load). Therefore, 1st September was the cut-off point and refractive error was the outcome. The difference in refractive error pre and post September 1 is a measure of the effect of education on refractive error. For each age group, RD was used to model the effect of discontinuity on refractive error (difference of mean RE and 95% CI) at the cut-off point. The RD model used non parametric local polynomial regression where weights for each data reduce as they move further from the cut-off point and the size of each bin to estimate the discontinuity effect is determined using mean square error.

2. When comparing the prevalence estimates of myopia in the discussion section, it is better to include a

figure. In addition, the authors might include more studies reporting the prevalence of myopia in Chinese children.
Response: Thanks. We've added two papers reporting prevalence of high myopia in Chinese children and included a figure presenting the prevalence estimates for more intuitive comparisons (Figure 5).

3.The authors mentioned information bias using questionnaires. It is better if they could explain if the bias is differential or non-differential and its potential impact on effect size of ORs.
Response: Thanks for your comments. The questionnaire included both quantitative and qualitative questions.
The query on time spent on indoor/outdoor work may result in non-differential bias as it is based on recall. The bias induced by recall may be avoided by collecting the data objectively using wearables – we have discussed this as a potential limitation.

The qualitative queries such as 'sitting too close to television' and 'holding a book too close while reading' are influenced by differential bias and therefore may impact on the effect size of ORs. However, an impact of differential bias on the OR is greater for the middle category ("sometimes") rather than extreme category ("usually") which myopes are likely to over emphasis compared to non-myopes. With the lower impact on extreme categories the overall inference of these factors are likely to remain robust, however we acknowledge that it is a potential limitation of use of questionnaires and have included it in the discussion.

Reviewer: 2
Dr. R Baraas, University of South-Eastern Norway
Comments to the Author:
I find the data presented in this manuscript very interesting. The dataset is large and there is no doubt that myopia is a major public health burden in Shanghai, both today and in the future. There are aspects that would benefit from more information and a more elaborate discussion, especially related to the causal effect of the school start date on the refractive error. The authors have not addressed important uncertainties that are needed to substantiate that the reported 0.67D difference over 10 years is "a striking effect of schooling/education resulting in a more myopic refractive error"?

We appreciate the insightful and positive comments from the reviewer.

Two steps have been made by the authors for ensuring a conservative prevalence estimate of myopia when non-cycloplegic data is the basis for the classification. First, myopia was defined as =< -1.00 D. Second, a model (Equation 1) was used to correct for known errors with classification of refractive error based on non-cycloplegic autorefraction. Why did the authors choose -1.00D as the alternative definition rather than, let's say, -0.75 or -1.25D? Was this related to the uncertainty with the model in reclassifying individuals or was it arbitrarily chosen to fit previously published prevalence estimates? The model reportedly only classifies 77% of the eyes correctly (Sankaridurg et al. 2017). What are the consequences that over 20% of eyes could have been misclassified, both in relation to prevalence, but also to the outcome of the regression discontinuity analysis?

As the data was non-cycloplegic to arrive at the best possible estimate of prevalence, we
a) Adopted the methodology as in the article Sankaridurg 2017 and used the equation as provided in the article. Indeed, as the reviewer rightly pointed out, the equation was found to classify 77% of the eyes.
b) Therefore, to improve the sensitivity, we used a higher cut-off value of -1.0D rather than -0.75D. As mentioned in the discussion, a previous publication from the same region used similar criteria of -1.00D

and cycloplegia- comparison indicates that our data is somewhat conservative (slightly less prevalence reported). Using a higher cut off criteria of -1.25D would have prevented us from comparing the results to other prevalence figures from the region.
c) Since the regression discontinuity analysis uses absolute refractive error rather than prevalence figures, we do not believe there would be an effect on the discontinuity analysis.

The model includes visual acuity measures. It is a strength that logMAR charts were employed, and that luminance was measured in the cabinets used, however, the range of luminance was considerable (80–320 cd/m2), could this have influenced the percentage of misclassifications as the model was based on visual acuity measures obtained at 400 cd/m2 (Sankaridurg et al. 2017)?

Since visual acuity was measured using high contrast letters we believe the influence of a decrement in luminance might not have been as substantial as compared to using lower contrast letters. However, assuming that luminance did influence visual acuity, poorer luminance would have resulted in poorer VA which then increased the risk of the eye being misclassified an eye as myopic. Therefore, using a higher myopia estimate than the -0.75D used by Sankaridurg et al. 2017 provides a benefit in reducing the risk of misclassification.

Regression discontinuity design (RDD) is based on a number of assumptions, and an assessment of whether these assumptions are met in this study needs to be included.
Agree

An advantage of RDD is its ability to estimate local causal effects in a small range near the cut-off, but there is no information on what the range was: was it all the children born in August versus all the children born in September?

We agree. The RD model was not for all children from all years. A separate model was developed for each age group (5 to 14 years), - for each year/age group, all children before August 31 were binned into pre-September category and all children born on or after 1 September were binned into post Sept category. To determine the effect of discontinuity, the RD model estimates the bin size based on mean square error. The bin size in days was 34, 49, 41, 39, 34, 30, 43, 30, 39 and 29 days for ages 5,6,7,8,9,10,11,12,13,14 years respectively.

There are discontinuities at locations other than the cut-off: might these temporal trends confound the results? Only means are presented in Figure 1, there are no actual data points.

We have revised Figure 1 to present the observed RE at each month along with its 95% error bar. Each series in the graph represents age in completed years and indicates that there is no significant variability at each time point other than the observed discontinuity at the cut-off date. With respect to the temporal trends, we are not entirely clear if the reviewer was commenting on the overall temporal trend across the ages – if so, this is accounted for by developing models for each age. On the other hand, if the reviewer is commenting on the temporal trend within each age - there are some variations observed especially for the higher grades but these appear to be much smaller than the discontinuity observed at the cut-off.
We have now added one more graph (Figure 2) that shows the scatter of the data for each age and polynomial line based on the RD model with reference to the cut-off date.

Revised Figure 1

New Figure 2
Neither are the actual regression lines with confidence intervals on each side of the identified threshold included (in Figure 1). Thus, the novel aspect in this manuscript, regression discontinuity (RD) analysis and the strength of the results from this analysis is difficult to evaluated properly.
The new figure 2 demonstrates the scatter of the data for each age group and polynomial line based on the local polynomial regression used in the RD model with reference to the cut-off date of 1st Sept. The graphs show a significant discontinuity at 1st Sep, where the intercept of the polynomial line after 1st Sep shows a lower RE. This was not observed in 5 year old kids. We are also reporting the bin sizes used by RD model for each age. Though pre-Sept accounted for 8 months and pPost Sept was only 4 months, the bin sizes used by the RD model were 34, 49, 41, 39, 34, 30, 43, 30, 37 and 29 days for ages 5,6,7,8,9,10,11,12,13,14 years respectively.

Error bars are given in Figure 2, but there is no information on whether these are SDs, SEs or CIs.
Figure 2 (is now figure 3) shows the RD model based estimates of the difference in RE due to the discontinuity. The error bars in Figure 3 is the 95% confidence interval. This is now mentioned in the manuscript.

Another issue is the asymmetry in number of participants in the pre- versus post- 1 September groups (339 963 in the pre group compared with 175 181 in the post group): are more children born pre- 1 September, and if so to what degree could this have affected RD analysis and results? Are there differences in socio-economic status, birth rates and frequency of caesarean section pre- and post- 1 September across the 17 districts in Shanghai, and if so to what degree could this have affected RD analysis and results? The uncertainties related to the above-mentioned issues have not been addressed by the authors.

We agree that the cut-off point of 1st Sep results in an asymmetry as there are more months prior to September (8 months) versus post September (4 months). The RD model uses local-polynomial estimator, which means data points closer to the cut-off point are weighted more than points further away from the cut-off, and therefore we believe that asymmetry would not affect the estimation substantially. Moreover, the bin sizes used by the RD model are not substantially different between ages and are all within 2 months (34, 49, 41, 39, 34, 30, 43, 30, 37 and 29 days for ages 5,6,7,8,9,10,11,12,13,14 years respectively).
The reviewer has a really good argument related to the differential rates pre and post Sept- unfortunately we have not gathered data for example on socio economic status. We observed that there a variation in birth rates by months (greater number of births for example in the months leading up to September- in August it was 12.2% versus 8.9% in September), however we do not fully understand impact of such factors and their dependencies- we therefore added this as a potential limitation. We have included this in the discussion.

Known issues related to data from questionnaires has been acknowledged by the authors.
Thank you

Other comments:
Objectives in abstract refers to "geographical variation", but there are no results related to the geographical variation across the 17 districts in Shanghai.
We apologise- this stemmed from a prior draft. We have deleted reference to “geographical variation” in

abstract.

There are no error bars in Figure 3.
We have now added 95% error bars but they are extremely small and difficult to see due to the large sample size.

Line 41: It is unclear what information was given to participants and if it was the participant or one or both parents/carers who gave consent.
Consent was obtained from at least one parent/carer. Parents were informed of the study prior to any examination. Participants were not provided information. Details of the process were explained in the methodology article published previously, where related supporting information has also been provided (He X, Zhao R, Sankaridurg P et al. Design and methodology of the Shanghai child and adolescent large scale eye study (SCALE). Clin exp ophthalmol, 46($), 329-338, 2018.). We've revised accordingly.

Was the questionnaire validated in some way? The questionnaire should be included as an appendix.
Yes- a previously validated questionnaire (T H, Y HQ, W L, Q X, L Y, L F. Development reliability and validity of child vision care related behavior assesment scale. Chin J Sch Health 2009; 30: 1007–10) was used. The entire visit record form including the questionnaire has been published as part of the study methodology. (He X, Zhao R, Sankaridurg P et al. Design and methodology of the Shanghai child and adolescent large scale eye study (SCALE). Clin exp ophthalmol, 46($), 329-338, 2018.)
Reviewer: 1
Competing interests of Reviewer: None declared
Reviewer: 2
Competing interests of Reviewer: None declared.

VERSION 2 – REVIEW

REVIEWER	Pan, Chen-wei Medical College of Soochow University, School of Public Health
REVIEW RETURNED	26-Jul-2021

GENERAL COMMENTS	The authors have addressed the comments.

REVIEWER	Baraas, R University of South-Eastern Norway, National Centre for Optics, Vision and Eye Care
REVIEW RETURNED	06-Sep-2021

GENERAL COMMENTS	I am happy with your response to my concerns and the revision to the manuscript.